Identifying the core bacterial microbiome of hydrocarbon degradation and a shift of dominant methanogenesis pathways in the oil and aqueous phases of petroleum reservoirs of different temperatures from China

Zhichao Zhou¹, Bo Liang², Li-Ying Wang², Jin-Feng Liu², Bo-Zhong Mu², Hojae Shim³, and Ji-Dong Gu¹

¹Laboratory of Environmental Microbiology and Toxicology, School of Biological Sciences, The University of Hong Kong, Pokfulam Road, Hong Kong, Hong Kong SAR, People’s Republic of China
²State Key Laboratory of Bioreactor Engineering and Institute of Applied Chemistry, East China University of Science and Technology, Shanghai 200237, People’s Republic of China
³Faculty of Science and Technology, University of Macau, Macau, People’s Republic of China

Correspondence: Ji-Dong Gu (jdgu@hku.hk) and Bo-Zhong Mu (bzm@ecust.edu.cn)

Received: 5 November 2018 – Discussion started: 26 February 2019
Revised: 20 September 2019 – Accepted: 7 October 2019 – Published: 8 November 2019

Abstract. Microorganisms in petroleum reservoirs play significant roles in hydrocarbon degradation, and through the terminal electron-accepting process of methanogenesis, they also contribute to microbially enhanced oil recovery (MEOR) worldwide, with great economic and environmental benefits. Here, a molecular investigation, using the 16S rRNA and mcrA gene profiles based on MiSeq sequencing and clone library construction methods, was conducted on oil and water (aqueous) phases of samples of high (82–88°C), moderate (45–63°C), and low temperatures (21–32°C) from seven petroleum reservoirs in China. A core bacterial microbiome with a small proportion of shared operational taxonomic unit (OTU) values, but a high proportion of sequences among all reservoirs was discovered, including aerobic degraders, sulfate- and nitrate-reducing bacteria, fermentative bacteria, and sulfur-oxidizing bacteria distributed mainly in Proteobacteria, Bacteroidetes, Deferribacteres, Deinococcus–Thermus, Firmicutes, Spirochaetes, and Thermotogae. Their prevalence in the previously reported petroleum reservoirs and successive enrichment cultures suggests their common roles and functions involved in aliphatic and aromatic hydrocarbon degradation. The methanogenic process generally shifts from the dominant hydrogenotrophic pathway in the aqueous phase to the acetoclastic pathway in the oil phase in high-temperature reservoirs, but the opposite was true for low-temperature reservoirs. No difference was detected between the two phases in moderate temperature reservoirs. Physicochemical factors, including pH; temperature; phase conditions; and nitrate, Mn²⁺, and Mg²⁺ concentrations were the main factors correlated to the microbial compositional and functional profiles significantly. Linear discriminant analysis (LDA) effect size (LEfSe) analysis shows distribution differences of microbial groups towards pH, temperature, and the oil and aqueous phases. Using the software Tax4Fun for functional profiling indicated functional metabolism differences between the two phases, including amino acids, hydrocarbons in the oil phase, and carbohydrates in the aqueous phase.

1 Introduction

Microbially enhanced oil recovery (MEOR) is one of the most feasible and profitable technologies for extracting residual oil from low-productivity reservoirs, by utilizing microorganisms as the major functional players for hydrocarbon transformation via fermentation and methanogenesis in MEOR applications (Magot et al., 2000; Mbadinga et al., 2011; Mesle et al., 2013). Implementation of MEOR could also attenuate side effects by injection of nitrate into oil wells to stimulate the propagation of nitrate-reducing bacte-
ria (NRB) and inhibit the growth of sulfate-reducing bacteria (SRB), subsequently reducing the negative effects by sulfido- 
genic activity on corrosion and oil quality deterioration (Gao et al., 2013). Currently, investigations of compositional and 
functional profiles of petroleum reservoir microbiomes have attracted attention from both scientists and engineers because 
their information will improve our understanding of micro-
bial diversity and function in oil reservoir systems and im-
plement MEOR through manipulating microbial community 
composition and activities (Magot et al., 2000; Mesle et al., 
2013; Youssif et al., 2009).

Petroleum reservoirs are often characterized as high-
temperature and high-pressure subterranean ecosystems, but 
low-temperature reservoirs also exist around the world 
(Grabowski et al., 2005; Li et al., 2013; Voordouw et al., 
1996; Li et al., 2014). The physicochemical and geographic 
differences among petroleum reservoirs, together with ex-
genous injection of microorganisms and nutrients during 
MEOR operation, could affect the composition and func-
tional groups of the indigenous microbial community. Many 
previously discovered bacterial and archaeal groups have 
important functional roles, e.g., sulfate-reducing, nitrate- 
or nitrite-reducing, and fermentative bacteria; thermophilic ar-
chaecal hydrocarbon degraders; and methanogens, as well as 
exogenous aerobic hydrocarbon degraders, in the microbial 
community (Gao et al., 2015a,b; Lenchi et al., 2013; Li et al., 
2012a; Orphan et al., 2003; Wang et al., 2012; Grabowski et al., 
2005; Pham et al., 2009). The common and specific roles 
and functions of these microorganisms in petroleum reser-
vours are of great value and should be comprehensively ex-
amined. Methanogens cooperate with hydrocarbon degraders and 
fermentative bacteria synergistically to make the overall 
degradation processes thermodynamically favorable under 
largely anaerobic conditions (Mbadinga et al., 2011). Fer-
mentative products from hydrocarbon degradation provide 
the essential growth substrates for methanogens with two 
major pathways, namely hydrogenotrophic and acetoclas-
tic methanogenesis (Magot et al., 2000). In many cases the 
syntrophic acetate oxidation pathway coupling with hy-
drogenotrophic methanogenesis is prevalent in both in situ 
and enrichment incubation ex situ (Mesle et al., 2013; Liang 
et al., 2015; Wang et al., 2012; Lee et al., 2015; Mayumi 
et al., 2011). However, very little is known of the factors 
that shape the methanogenic pathways and the structure of 
methanogenic communities.

Petroleum reservoir is a complex system consisting of 
porous sandstones with oil, water, and air. Microorganisms 
attached to the oil phase of petroleum fluids are largely 
neglected in majority of the previous investigations due 
to technical difficulties in DNA extraction and sequencing 
(Kryachko et al., 2012). Oil-attached microorganisms in-
fluence the oil–water interface properties via the produc-
tion of biosurfactants and metabolites or the formation of 
adsorbed colloidal particles to enhance oil recovery per-
formance (Kobayashi et al., 2012; Kryachko et al., 2012; 
Wang et al., 2014). The distribution, function, and contribu-
tion of these microorganisms to MEOR success could be 
very different from those of the aqueous phase (Kobayashi 
et al., 2012). Oil-degrading microorganisms, including ther-
mophilic hydrocarbon-degrading archaea and methanogens, 
play different roles in the MEOR process (Mbadinga et al., 
2011). Their compositional patterns and functional profiles 
in terms of temperature and the oil and aqueous phases are 
of great value for the understanding of the mechanism of 
MEOR. Knowledge of the major microbial drivers, their po-
tential functions, distribution characteristics, and changing 
patterns towards environmental parameters should be one 
of the research directions taken for a better understanding 
of the MEOR process. In this study, the research objectives 
were to reveal the compositional and functional differences 
of petroleum-reservoir-inhabiting microorganisms under dif-
ferent temperatures (high: 82–88 °C; moderate: 45–63 °C; 
and low: 21–32 °C), the methanogenesis pathways in the oil 
and aqueous phases of these samples, and the influence of 
physicochemical factors on microbial community compo-
nition.

2 Materials and methods

2.1 Characterization of geographic properties of 
sampling reservoirs

Petroleum production fluid samples were collected from 
seven oilfields across China, covering oil wells of differ-
ent geographical locations and temperatures. The reservoirs 
and crude oil properties, together with the the aqueous-phase 
chemical characteristics of this study, are listed in Table 1. 
Detailed reservoir properties are described in the Supple-
ment.

2.2 Sample collection, ion concentration measuring, 
and DNA extraction

Each sample, containing a mixture of crude oil and water, 
was collected on site in a sterilized container after flush-
ing each wellhead for at least 3–5 min. The containers were 
screw capped to avoid air invasion and transported back to 
the nearby laboratory immediately for further processing. 
The oil and water mixture was gently heated to 50 °C to make 
it semifluid and then separated into oil and aqueous phases 
in a separatory funnel. Heating was operated as quickly as 
possible according to the solidification degree of individ-
ual petroleum fluid samples, in order to reduce the lytic ef-
fect of microbial cells within them. Ion concentrations of 
the aqueous-phase samples were measured using Dionex 600 
ion chromatography (Triad Scientific, Inc., Manasquan, NJ, 
USA) following the manufacturer’s instructions.

To obtain aqueous-phase DNA, the aqueous phase of 
each sample after oil and water separation was first filtered
Table 1. Reservoir geophysical properties and physicochemical properties of aqueous-phase samples.

<table>
<thead>
<tr>
<th>Location name</th>
<th>Shengli-Zhan3-26</th>
<th>Huabei-B51 (Karamay)-Liuzhong</th>
<th>Xinjiang (Karamay)-Qixi</th>
<th>Xinjiang-Wei5</th>
<th>Jiangsu-Gao6</th>
<th>Jiangsu-Qinying</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample name</td>
<td>P1</td>
<td>P2</td>
<td>P3</td>
<td>P4</td>
<td>P5</td>
<td>P6</td>
</tr>
<tr>
<td>Depth (m)</td>
<td>~ 1300</td>
<td>~ 1101</td>
<td>~ 480</td>
<td>~ 1088</td>
<td>1018</td>
<td>1970</td>
</tr>
<tr>
<td>Temp (°C)</td>
<td>63</td>
<td>45</td>
<td>21</td>
<td>32</td>
<td>53.5</td>
<td>82</td>
</tr>
<tr>
<td>pH</td>
<td>7.1</td>
<td>7.2</td>
<td>7.0</td>
<td>7.1</td>
<td>8.62</td>
<td>8.88</td>
</tr>
<tr>
<td>Effective porosity (%)</td>
<td>30</td>
<td>22.2</td>
<td>20.5</td>
<td>17.4</td>
<td>21.7</td>
<td>17.4</td>
</tr>
<tr>
<td>Average permeability (\times 10^{-3} \mu m^2)</td>
<td>800</td>
<td>12.6</td>
<td>466</td>
<td>274</td>
<td>91.8</td>
<td>94.5</td>
</tr>
<tr>
<td>Oil viscosity (mPa s)</td>
<td>1720</td>
<td>402</td>
<td>417</td>
<td>44.8</td>
<td>292</td>
<td>8.4–83.9</td>
</tr>
<tr>
<td>Water flooding operation (years)</td>
<td>22</td>
<td>4</td>
<td>38</td>
<td>46</td>
<td>13</td>
<td>18</td>
</tr>
</tbody>
</table>

| Cl\(^-\) (mg L\(^{-1}\))    | 3850              | 819                           | 3864                     | 2000         | 7563.01     | 1155.95         |
| SO\(_4^{2-}\) (mg L\(^{-1}\)) | 2244             | 32.4                          | 124.8                    | 7.7          | 1921.06     | 156.81          |
| PO\(_4^{3-}\) (mg L\(^{-1}\)) | 0.1               | Nd                            | Nd                       | Nd           | 139.77      | 6.29            |
| NO\(_3^{-}\) (mg L\(^{-1}\)) | Nd                | Nd                            | 34.1                     | 1.4          | Nd          | 0.44            |
| Na\(^+\) (mg L\(^{-1}\))    | 3313              | 1064                          | 4196                     | 5399         | 7363.52     | 1593.55         |
| K\(^+\) (mg L\(^{-1}\))     | 94.2              | 22.3                          | 35.1                     | 45.6         | 9.36        | 3.26            |
| Ca\(^{2+}\) (mg L\(^{-1}\)) | 195.6             | 53.0                          | 103.3                    | 128.2        | 17.42       | 7.24            |
| Mg\(^{2+}\) (mg L\(^{-1}\)) | 46.1              | 17.6                          | 44.7                     | 64.0         | 10.27       | 2.25            |
| Mn\(^{2+}\) (mg L\(^{-1}\)) | 0.3               | 0.1                           | 0.3                      | 0.4          | Nd          | Nd              |
| Acetate (mg L\(^{-1}\))     | 32                | 57.9                          | 344                      | 6.97         | 44.07       | 285.66          |

“Nd” means not detected. No aqueous phase was obtained from Jiangsu-Qinying reservoir sample.

2.3 Clone library construction and analysis

The methyl-coenzyme-M reductase-encoding gene \(mcrA\) primer pair, ME3MF and ME3MF-e–ME2r′ (Narihiro and Sekiguchi, 2011), and the 16S rRNA gene primer pair targeting methanogens, MetA86f–ARC915r (Wright and Pimm, 2003; Narihiro and Sekiguchi, 2011; Yu et al., 2008), were used to generate gene clone libraries for each sample (aqueous and oil phases separately) according to the method established previously (Zhou et al., 2015) (Table 2).

A FunGene pipeline chimera check was applied to check \(mcrA\) gene sequences using the UCHIME de novo mode (Edgar et al., 2011). USEARCH software was used to check chimeras of methanogenic 16S rRNA gene sequences using a QIIME-compatible SILVA 119 release SSURef database (“rdp_gold fasta”) file as a reference. Then, for \(mcrA\) gene sequences, a de novo operational taxonomic unit (OTU) picking method was applied by QIIME at a cutoff value of 0.05 (Caporaso et al., 2010). Representative OTU sequences were aligned and inserted into the \(mcrA\) gene ARB project database through the maximum parsimony method without changing the initial tree topology (Angel et al., 2012; Ludwig et al., 2004). The phylogenetic affiliation was assigned, and taxonomic composition results were processed by QIIME accordingly (Fig. S1 in the Supplement). For methanogenic 16S rRNA gene sequence clustering and diversity analysis, the same method was applied via QIIME as described in the followings.

2.4 MiSeq sequencing and QIIME-based analysis

The prokaryotic universal primer pair 515F–909R (Caporaso et al., 2012; Wang and Qian, 2009) and archaeal universal primer pair Arch347F–Arch806R (Takai and Horikoshi, 2000) were used to amplify samples of this study (both with barcodes attached to the forward primers, Table 2). Two independent polymerase chain reaction (PCR) reactions were conducted and then combined to yield enough PCR products to compromise variations between different batches. Then, pooled PCR products of each sample with approximately 100 or 200 ng DNA were subjected to one MiSeq run. Sequencing samples were prepared using TruSeq DNA Kit according to the manufacturer’s instruction. The library was uploaded to an Illumina MiSeq platform for sequencing with reagent...
Table 2. PCR primers and PCR settings.

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequences</th>
<th>Primer final concentration</th>
<th>Template DNA quantity</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>mcrA gene primer</td>
<td>ME3MF and ME3MF-e</td>
<td></td>
<td>1 µM</td>
<td>10–20 ng</td>
</tr>
<tr>
<td></td>
<td>ME3MF: ATGTCNGGTGGHHTUMGGSTTYAC;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ME3MF-e: ATGAGCGGTGGTGTGCTGGTTAC;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>concentration ratio = 250 : 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ME2′</td>
<td></td>
<td>1 µM</td>
<td>10–20 ng</td>
</tr>
<tr>
<td></td>
<td>TCACTBGCGATTDGGCRTAGT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methanogenic 16S rRNA</td>
<td>MetA86f</td>
<td>300 nM</td>
<td>10–20 ng</td>
<td></td>
</tr>
<tr>
<td>gene primer</td>
<td>GCTCAGTAACACGTGG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>95 °C for 2 min; 35 × (95 °C for 30 s, 59.5 °C for 30 s and 72 °C for 40 s); 72 °C for 5 min, and 4 °C for 2 min.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prokaryotic universal</td>
<td>515F (12 nt barcode added)</td>
<td>1 µM</td>
<td>10 ng</td>
<td></td>
</tr>
<tr>
<td>primer</td>
<td>GTCGCCAGCMGCCGCGGTAA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>94 °C for 3 min; 30 × (94 °C for 40 s, 56 °C for 60 s and 72 °C for 60 s); 72 °C for 10 min, and 4 °C for 2 min.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Archaeal universal</td>
<td>Arch347F (12 nt barcode added)</td>
<td>1 µM</td>
<td>10 ng</td>
<td></td>
</tr>
<tr>
<td>primer</td>
<td>GYGCCAGCGCGCGGTAAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>95 °C for 5 min; 32 × (95 °C for 30 s, 50 °C for 30 s and 72 °C for 40 s); 72 °C for 10 min, and 4 °C for 2 min.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After merging paired-end reads from raw sequencing data with FLASH-1.2.8, the fastx toolkit was applied to split merged reads from one run into individual samples according to the primer barcodes (Magoc and Salzberg, 2011). Then, all sequences were split into each library with the name of each sample attached according to the barcode map using the QIIME command “split_libraries” (Caporaso et al., 2010). The criterion for filtering out underqualified sequences was “s 15 -k -a 6 -r -l 150 -b 12 -M 5 -e 0”. Chimera checking was conducted by USEARCH software using the QIIME-compatible SILVA 119 release SSURef database (“rdp_gold” fasta) file as the reference (Edgar, 2010). Clustering; picking OTU; taxonomy assignment, aligning, filtering alignments, and phylogenetic tree construction; taxonomic composition summarizing; and alpha and beta diversity analyses were conducted step-by-step by the QIIME pipeline with QIIME-compatible SILVA 119 SSURef database as the reference (Caporaso et al., 2010). In clustering, the “pick_open_reference_otus.py” command was used to conduct OTU dividing, and the BLAST method was used to assign taxonomy to input sequences. This subsampled open-reference OTU picking method was the performed optimized and optimal strategy suggested by the developers (Rideout et al., 2014). After OTU table “biom” files were generated, in order to get the bacterial community composition information from prokaryotic 16S rRNA gene primer amplified libraries, the “filter_taxa_from_otu_table.py” command was introduced to only retain bacterial OTUs in the “biom” file. Similarly, exclusive archaeal and methanogenic OTU table files could also be processed from archaeal 16S rRNA gene primer amplified libraries. Summary information of each sample OTU abundance could be calculated by the “biom summarize-table” command, and then the lowest number among all samples was chosen as the subsampling size to make each library acquire an even size using the “multiple_rarefaction_even_depth.py” command. The taxonomic compositional table was drawn according to the subsampled biom file. Since there was no lane mask file available in this SILVA-compatible 119 release SSURef database, an alignment filtering method was performed independently with an entropy threshold of 0.1 and gap filter threshold of 0.9 after obtaining aligned sequences via the PyNAST method. Diversity parameters of each library could be generated by alpha diversity calling, with the rarefaction curve; Good’s coverage value; and Shannon, Chao1, Simpson, and phylogenetic diversity (PD) whole tree indices calculated. Beta diversity, which delineates the dissimilarity relationship among samples, was generated and visualized through unweighted and weighted UniFrac matrix and non-phylogenetic Bray–Curtis matrix method. The
pairwise-shared OTU numbers were calculated from “biom” by the command “shared_phylogenotypes.py”. Core microbiome (shared OTU table in a specific sample category) was identified by the “compute_core_microbiome.py” command.

2.5 Diversity and statistical analysis

The statistical significance of the community composition of samples among different categories was valued by the anosim and adonis methods, implemented in “compare_categories” command in QIIME. A Mantel test was used to compare the distance matrix of physicochemical parameters and UniFrac and Bray–Curtis distance matrix from beta diversity analysis by QIIME. For aqueous-phase samples, both of the in situ physicochemical parameters and ion concentrations were used in the analysis, while only in situ physicochemical parameters were included for oil-phase samples. The compositional bar chart and bubble chart were modified and illustrated from taxonomic summary results. Tax4Fun was used to predict the functional capabilities based on abundance profiles of microbial 16S rRNA gene datasets (Asshauer et al., 2015). Linear discriminant analysis (LDA) effect size (LEfSe) analysis was applied to illustrate the biomarker species with high statistical significance in different sample categories and the functional profiles statistically distributed in different sample categories (Segata et al., 2011).

2.6 Quantitative PCR on mcrA gene abundance

The quantitative PCR measurement was conducted using the iTaq™ Universal SYBR® Green Supermix Kit (BIO-RAD). The qPCR mixture contained in 15 µL: 7.5 µL of Supermix, 16 µg of bovine serum albumin (BSA) (Roche), and 1 µM final concentration of the primer pair, i.e., ME3MF and ME3MF-e (250 : 1) and ME2r. Annealing temperature was set to be the same as the clone library PCR setting, and the rest thermocycling settings were according to the manufacturer’s instructions. A single randomly picked pMD18-T plasmid with the mcrA gene inserted was used to make the standard curve. The DNA template concentration was adjusted to 0–40 ng µL⁻¹. Results that deviated significantly from values in the replicate groups were omitted, and undetermined results (under the detection limit) were also deleted. The property of the final adjusted standard curve is \( r^2 = 0.995 \) and Eff% = 83.32.

2.7 Sequencing result deposition

For clone library sequences, KT314862–KT315353 were assigned to methanogenic 16S rRNA gene sequences and KT314340–KT314835 were assigned to mcrA gene sequences. For high-throughput sequencing data based on Illumina MiSeq platform, PRJEB9855 was assigned to prokaryotic 16S rRNA gene sequencing data and PRJEB10996 was assigned to archaeal 16S rRNA gene sequencing data.

3 Results and discussion

3.1 Common OTU among different categories and core bacterial microbiome

Community composition results showed that 21 bacterial phyla were obtained with an average abundance of more than 0.1 % (Fig. 1a) and that the three major archaeal phyla were Thaumarchaeota, Euryarchaeota, and Crenarchaeota (Fig. 1b). Pairwise-shared OTU numbers of all samples indicated that, irrespective of combinations between aqueous- and oil-phase samples, the average numbers of shared bacterial OTUs ranged from 199.9 to 292.4, accounting for 26.6 %–36.2 % of the total OTU numbers within individual samples; average numbers of shared archaeal OTUs ranged from 1.8 to 11.9, accounting for 8.5 %–23.4 % of total OTU numbers within individual samples (Table S4 in the Supplement). Core bacterial OTU numbers among aqueous, oil, and all samples were 73, 57, and 46, which accounted for 7.1 %–10.1 %, 5.9 %–10.3 %, and 4.5 %–8.3 % of OTU numbers in individual samples. The core archaeal OTU number among aqueous samples was only 3, accounting for 3.9 %–8.1 % of OTU numbers in individual samples, and no archaeal OTU was shared among oil-phase samples.

However, by investigating taxonomic profiles of core bacterial OTUs, the shared OTUs were 49, 41, and 34 genera in aqueous, oil, and all samples, corresponding to 65.5 %, 59.9 %, and 58.8 % of average sequences in the total bacterial community, respectively (Tables S5 and S6). Most of the core bacterial OTUs belonged to the most abundant 36 genera, of which the numbers of shared genera among aqueous, oil, and all samples were 28, 23, and 23, respectively (Fig. 2, and Tables S5 and S6).

There was no significant difference of shared bacterial OTU numbers within and between aqueous- and oil-phase samples, suggesting a core microbiome was shared among all components. The core OTUs covered around two-thirds of the total bacterial sequences, even though the percentages of core or total OTU number for individual samples were 4.5 % to 10.3 %. The core microbiome shared among all petroleum reservoirs could be the key participants mediating critical microbial processes, such as activation, degradation, fermentation, oil emulsification, and methane generation (Yamane et al., 2008; Wang et al., 2014; Pham et al., 2009; Orphan et al., 2000; Magot et al., 2000). This spectrum of core microbiome shares common functional roles in facilitating MEOR and is modified by the in situ physicochemical conditions of different reservoirs (Fig. 2 and Table S5). It is important to connect the major microbial players, including their community compositions and specific functional capacities, to the interpretation of MEOR processes in the petroleum reservoirs. Meanwhile, the core microbiome serves as a good basis for simplifying microbial participants and, primarily, their roles in the petroleum reservoirs and is useful for modeling and monitoring the MEOR processes of petroleum reservoirs from dif-
different locations. Moreover, substantial portions of the aerobic bacteria being discovered in the core bacterial microbiome across different reservoirs imply that exogenous bacteria introduced into subsurface reservoirs by water flushing can be also represented in the core composition and it plays important roles in enhancing oil recovery (Gao et al., 2015b; Youssef et al., 2009).

Bacterial community distribution in the aqueous phase showed correlation with temperature via the Mantel test using unweighted UniFrac matrix method (Table S8). Additionally, temperature also significantly affected the dissimilar distance matrix of PCoA coordinates of the bacterial community via both adonis and anosim statistical analyses (Tables S3 and S7). However, temperature changed the proportions of the taxa of the core microbiome significantly but did not change their presence or absence (Tables S3, S7 and S8). Consequently, the common genera discovered from this study show that, despite geographic and physicochemical differences, there is a core microbiome with small OTU numbers but large sequence proportions in these petroleum reservoirs across China that are possibly responsible for mediating hydrocarbon degradation processes. Prevalent core genera discovered in this study also appear to be the most abundant components, as previously reported in petroleum reservoirs within and beyond China (Gao et al., 2015a, b; Li et al., 2012a; Orphan et al., 2003, 2000; Tang et al., 2012; Yamane et al., 2008; Dahle et al., 2008; Li et al., 2015, 2014; Liang et al., 2015) and enrichment cultures from petroleum reservoirs worldwide (Gray et al., 2011; Li et al., 2012b; Liang et al., 2015; Wang et al., 2011).

3.2 Shift of major methanogenesis pathways between the oil and aqueous phases

The methanogenic process generally shifts from the dominant hydrogenotrophic pathway in the aqueous phase to the acetoclastic pathway in the oil phase of high-temperature reservoirs. The opposite is true for low-temperature samples, but no difference was detected between the two phases in moderate temperature reservoirs (Fig. 3). The different patterns of shifting of the major methanogenesis mechanisms in the aqueous and oil phases were evident in different temperature petroleum reservoirs of this study. Quantitative measurements of mcrA gene in both the aqueous and oil phases of all samples are summarized in the Supplement (Fig. S1).

The underlying methanogenesis mechanism could change substantially in reservoirs with distinctive geochemical prop-
Figure 2. Average abundances and functional roles of core bacterial microbiome in all petroleum samples (including aqueous- and oil-phase samples). Core microbial OTUs (23 OTUs) were affiliated to the most abundant genera. The functional roles were assigned according to reference publications (Supplement). Degradation pathway was modified from Mbadinga et al. (2011). Abbreviations: Alpha., Beta., Gamma., Delta., Epsilon. (Alpha-, Beta-, Gamma-, Delta-, Epsilonproteobacteria), D.-T. (Deinococcus–Thermus), Bac. (Bacteroidetes), Deferri. (Deferribacteres), Thermo. (Thermotogae), Spiro. (Spirochaetes), and Fir. (Firmicutes). UG stands for uncultured genus.

Figure 3. Bubble chart of the proportional composition of methanogens based on MiSeq sequencing results of archaeal 16S rRNA genes and clone sequencing results of the mcrA gene and methanogenic 16S rRNA gene. The major substrate utilization properties originated in Liu and Whitman (2008). MeNH₂ is methylamine and substrates in parentheses refer to being utilized by some but not all species. Methanothermobacter thermautotrophicus strain Delta H is the phylogenetically closest cultured clone to Methanobacteriaceae archaeon 15aZ. Thus, the combination of Methanothermobacter and the Methanobacteriaceae archaeon 15aZ was used. Methanogenesis shift was based on the transition of major methanogenesis pathway. Abbreviations: Aceto. M., Hydro. M., and Methylo. M. are Acetoclastic methanogenesis, Hydrogenotrophic methanogenesis, and Methylotrophic methanogenesis, respectively.
erties (Magot et al., 2000). The dominance of methylotrophic methanogenesis is rarely observed in petroleum reservoirs. It is claimed that there is a very low possibility that methyl-compounds could be generated during the degradation of kerogen (Mesle et al., 2013). However, the dominance of obligate methylotrophic Methanolobus was found in all water, rock and coal samples in a coal bed methane reservoir within the eastern Ordos Basin of China (Guo et al., 2012). It could be deduced that the prevalence of methylotrophic methanogenesis under certain conditions is directly fueled by the existence and availability of methyl-containing substances. Since the relative abundances of bacteria in petroleum reservoir samples are always higher than those of methanogenic community, the methanogenesis process might not be the dominant process among all microbial processes. For instance, in P1 and P5 aqueous samples, sulfate concentrations are considerably high (Table 1) and geochemical conditions are more favorable for sulfate reduction than methanogenesis. A large proportion of Firmicutes and Deferribacteres in P5A were potentially responsible for the activities of sulfate reduction, according to their relative abundance (Fig. 1) and functional capacities (Fig. 2). It is suggested the methylated compounds could be produced by the degradation of glycine betaine (an osmoprotective agent) that are accumulated or generated by halophiles in saline petroleum (Olivier and Alazard, 2010). Subsequently, it fuels the growth of methylotrophic methanogens in certain petroleum environments. Although no molecular or chemical evidence was reported for this process in petroleum environments, a pioneer study on hydraulic fracturing in shales has detected glycine betaine as the major known osmoprotectant in the produced fluids, and the combined metagenomic and metabolite findings suggest the similar glycine betaine utilization pathways in fractured shales (Daly et al., 2016).

Numerous studies based on in situ or in enrichment incubation microcosms on the aqueous phase of reservoir fluids indicate that syntrophic acetate oxidation associated with hydrogenotrophic methanogenesis is the major hydrocarbon degradation pathway (Wang et al., 2012; Mayumi et al., 2011; Lee et al., 2015; Mbadinga et al., 2012; Gray et al., 2011). The dominance of hydrogenotrophic methanogenesis in subsurface ecosystems could result from the external hydrogen originating from maturation of organic matter and/or mineral hydrolysis (Head et al., 2003) and the synergistic effect, in association with acetate oxidizers, whereby acetate was firstly oxidized to H2 and CO2 is then utilized by methanogenesis (Liu and Whitman, 2008). Additionally, a stable isotope labeling experiment on oil-degrading microcosms showed that despite the coexistence of acetoclastic methanogenesis and acetate syntrophic oxidation in the initial state, the latter process prevailed over the former one when introducing a low acetate concentration initially (Gray et al., 2011). The above evidence suggests that acetate syntrophic oxidation could exceed acetoclastic methanogenesis and contribute substrates H2 and CO2 to potentially favor hydrogenotrophic methanogenesis process.

It is still difficult to determine whether the temperature has directly or indirectly been involved in the alteration of the methanogenesis pathways. Reservoir fluid constituents may affect methanogenic degradation because crude oil and creosote inhibit acetoclastic methanogenesis (Warren et al., 2004) and volatile hydrocarbons (nC5–nC10) inhibit methanogenic degradation rates without changing the abundances of both hydrogenotrophic and acetoclastic methanogens (Sherry et al., 2014). Since most of the currently available community data are based on the microbial assemblages within injection or production water, a new understanding of the local microbiome distribution and changes should focus on the oil or hydrophobic fraction of the reservoir fluids (Kobayashi et al., 2012; Tang et al., 2012; Lenchi et al., 2013). Consequently, the alteration pattern of major methanogenesis in the aqueous and oil phases under different temperature conditions could be further delineated. A combination of methods, including synthesis and quantification of the degradation intermediate (Bian et al., 2015), stable isotope labeling on tracing substrate transformation (Gray et al., 2011), and molecular analysis of the metabolically active microorganisms can advance the information of anaerobic degradation and methanogenesis processes in reservoir systems.

3.3 Physicochemical influence and taxa and function profiles

Temperature is an important physical factor shaping the community structure of bacterial (anosim and adonis p < 0.01) and methanogenic communities (anosim p < 0.05 and adonis p < 0.01) of the samples in this study (Fig. 4 and Table S7). Furthermore, a significant difference of taxa abundance among the three temperature categories for both bacterial and archaeal communities was evident by LEfSe analysis (Fig. 5). For the bacterial community, sample group (aqueous or oil phase from the same sample group), temperature (anosim and adonis p < 0.01), and pH (adonis p < 0.05) showed significant effects on separating samples into different categories. For the archaeal community, significant differences among sample categories were detected with sample group and temperature (both adonis p < 0.05) and pH and phase (anosim and adonis p < 0.05). For the methanogenic community, significant differences among sample categories were detected with sample group and temperature (sample group: anosim and adonis p < 0.05; temperature: anosim p < 0.05 and adonis p < 0.01) (Table S7).

The nitrate concentration dissimilarity matrix was significantly associated with all unweighted and weighted UniFrac and Bray–Curtis matrices (all p < 0.05) for aqueous bacterial community (Table S8). Meanwhile, methanogenic community in the oil phase was significantly affected by reservoir depth, temperature, pH, and water flooding operation years.
Figure 4. Principal coordinate analysis plot figures based on unweighted UniFrac matrices. Bacterial (a) and methanogenic (b) communities of 14 samples were separately analyzed to delineate the dissimilarity distances between each sample based on phylogenetic classification. The sample dots were categorized in terms of temperature (dot shape) and pH condition (color).

Figure 5. Cladogram based on LEfSe analysis results of bacteria (a, c) and archaea (b, d) in terms of temperature (a, b) and pH (c, d) categories. The taxonomic trees were generated from phylum to genus (inside to outside) in the hierarchical structure. Biomarker taxonomic levels were labeled as the color that had a logarithmic LDA score of at least 3.5. Pre-sample normalization was used to format the relative abundance data. All-against-all strategy was used in the multi-class analysis step.
based on association analysis of the weighted UniFrac matrix but not the unweighted UniFrac matrix (all \( p < 0.05 \) (Table S8), indicating that it was the abundance difference of certain taxa affecting the compositional pattern. More detailed relationships between physicochemical factors and bacterial, archaeal, and methanogenic communities are summarized in the Supplement (Tables S7 and S8).

The nitrate dissimilarity matrix was significantly correlated with dissimilarity matrices of all aqueous-phase bacterial communities using both unweighted and weighted UniFrac and Bray–Curtis matrices (all \( p < 0.05 \) (Table S8). Nitrate is an important chemical used in injection water to inhibit corrosion and maintain crude oil quality (Gao et al., 2013). Nitrate stimulates the growth of nitrate-reducing bacteria and inhibits the growth of SRB (Nemati et al., 2001; Gao et al., 2013). Consequently, nitrate injection shapes the microbial communities in petroleum reservoirs. \( \text{Mn}^{2+} \) and \( \text{Mg}^{2+} \) were shown to be strongly associated with the bacterial community of the aqueous phase based on the Bray–Curtis matrix (both \( p < 0.05 \) (Table S8). Metal ions can be electron acceptors for direct or indirect hydrocarbon degradation under anaerobic conditions (Mbadenga et al., 2011). Metal reducers could also utilize electrons from syntrophic partners to further facilitate direct aromatic hydrocarbon degradation (Kunapuli et al., 2007).

Differences in taxa and function profiles in the oil and aqueous phases were analyzed based on LEfSe and Tax4Fun (Figs. S2 and S3, and Table S9). For bacterial communities, amino-acid metabolism and xenobiotics biodegradation and metabolism were distributed more in the oil phase, while carbohydrate metabolism was distributed more in the aqueous phase. For archaeal communities, ubiquinone and other terpenoid quinone biosynthesis and butanoate and tryptophan metabolism, together with other glycan degradation pathways were distributed more in the oil phase. Since FTU (fraction of taxa that could be mapped to existing KEGG pathway) values of archaeal communities were unevenly distributed from sample to sample, the reliability of these functional predictions is in question (Table S9). Meanwhile, the database of Tax4Fun is far from completeness, due to the enormous amount of uncultured microorganism and their unknown genomes. To date, there are still very limited studies investigating the inhabiting preference of microbiomes in petroleum reservoirs. Meanwhile, their functional contributions to hydrocarbon degradation and methanogenesis in both the aqueous and oil phases remain elusive, which calls for further efforts towards this point (Kryachko et al., 2012; Wang et al., 2014; Kobayashi et al., 2012).

4 Conclusions

A core bacterial microbiome containing a small proportion of OTUs but a relatively large proportion of sequences mediating hydrocarbon degradation and fermentation was revealed by analysis of oil reservoirs of different temperatures. The core and common bacterial microbiome of major biodegrading functions were shared among geographically and physicochemically different reservoirs. The different and shifting patterns of the dominant methanogenesis pathway in the aqueous and oil phases within samples of different temperatures were evident. Factors of \( \text{pH} \); temperature; phase conditions; and nitrate, \( \text{Mn}^{2+} \), and \( \text{Mg}^{2+} \) concentrations shaped the microbial compositional and functional profiles significantly. Moreover, biomarker groups of bacteria and archaea associated with different \( \text{pH} \), temperature, and phase conditions indicate major differences in the biochemical function of amino acid metabolism, xenobiotics metabolism enriched in the oil phase, and carbohydrate metabolism enriched in the aqueous phase.

Data availability. The petroleum reservoir MiSeq 16S rRNA gene sequencing original datasets are available on the following GitHub page: https://github.com/ChaoLab/PetroleumMicrobiome (last access: 6 November 2019).

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/bg-16-4229-2019-supplement.

Author contributions. ZZ, BL, LYW, BZM, HS, and JDG conceived the project and designed the experiments. ZZ, BL, and LYW conducted the sampling, chemical, and molecular experiments. JFL, BZM, and JDG managed sample collection and supervised data interpretation. ZZ performed the original data analysis and drafted the original manuscript. All members contributed to refining the manuscript and approved the final version.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Xiangzhen Li’s group at the Chengdu Institute of Biology, Chinese Academy of Sciences, is thanked for their MiSeq sequencing efforts and related technical support. Kelly Lau is thanked for her supportive technician work. We are grateful for the support from local administrative and technical staff at the Shengli, Daqing, Huabei, Xinjiang Karamay, and Jiangsu oilfields.

Financial support. This research has been supported by the NSFC/RGC (grant no. N HKU718/11), the NSFC (grant nos. 41373070 and 41530318) and a postgraduate scholarship from HKU.

Review statement. This paper was edited by Tina Treude and reviewed by two anonymous referees.
References


